

Modeling hierarchical versus random exposure schedules in Pavlovian fear extinction: No
evidence for differential fear outcomes

Sara Scheveneels*, Yannick Boddez, Bram Vervliet, Dirk Hermans

Author Note

Sara Scheveneels, Center for the Psychology of Learning and Experimental
Psychopathology, KU Leuven, Leuven, Belgium.

Yannick Boddez, Department of Clinical Psychology and Experimental
Psychopathology, University of Groningen, Groningen, The Netherlands.

Bram Vervliet, Laboratory of Biological Psychology, KU Leuven, Leuven, Belgium.

Dirk Hermans, Center for the Psychology of Learning and Experimental
Psychopathology, KU Leuven, Leuven, Belgium

This work was supported by KU Leuven Program Funding Grant PF/10/005 awarded
to Dirk Hermans. We thank Maarten Vanherle for his assistance in collecting the data.

*Corresponding author:

Sara Scheveneels

Centre for Learning Psychology and Experimental Psychopathology, University of Leuven

Tiensestraat 102 – bus 3712

3000 Leuven, Belgium

Email: Sara.scheveneels@kuleuven.be

Abstract

In exposure therapy, the client can either be confronted with the fear-eliciting situations in a hierarchical way or in a random way. In the current study we developed a procedure to investigate the effects of hierarchical versus random exposure on long-term fear responding in the laboratory. Using a fear conditioning procedure, one stimulus (CS+) was paired with an electric shock (US), whereas another stimulus was not paired with the shock (CS-). The next day, participants underwent extinction training including presentations of the CS-, CS+ and a series of morphed stimuli between the CS- and CS+. In the hierarchical extinction condition (HE; $N = 32$), participants were first presented with the CS-, subsequently with the morph most similar to the CS-, then with the morph most similar to that one, and so forth, until reaching the CS+. In the random extinction condition (RE; $N = 32$), the same stimuli were presented but in a random order. Fear responding to the CS+, CS- and a new generalization stimulus (GS) was measured on the third day. Higher expectancy violation ($t(62) = -2.67, p = .01$), physiological arousal ($t(62) = -2.08, p = .04$) and variability in US-expectancy ratings ($t(62) = -2.25, p = .03$) were observed in the RE condition compared to the HE condition, suggesting the validity of this novel procedure. However, no differences between the RE and HE condition were found in fear responding as tested one day later, $F(1, 62) < 1$. In conclusion, we did not find evidence for differential long-term fear responding in modeling hierarchical versus random exposure in Pavlovian fear extinction.

Keywords: exposure therapy; extinction; exposure hierarchy; conditioning

Modeling Hierarchical versus Random Exposure Schedules in Pavlovian Fear Extinction: No Evidence for Differential Fear Outcomes

Exposure is a key component in the treatment of anxiety and related disorders (e.g., Choy, Fyer, & Lipsitz, 2007; Sánchez-Meca, Rosa-Alcázar, Marín-Martínez, & Gómez-Conesa, 2010). In exposure therapy, the patient is confronted with fear-eliciting stimuli or situations. In guiding therapists how to conduct exposure therapy, traditional models of exposure such as emotional processing theory (EPT) encourage the use of exposure hierarchies, whereby this confrontation takes place in a progressive manner (Foa, Huppert, & Cahill, 2006; Foa & Kozak, 1986). Typically, it is recommended to start the exposure with stimuli that elicit only little fear and to gradually progress towards the more distressing items. For example, a patient with fear of heights can first be exposed to standing on a 1st floor balcony with safety rails, moving stepwise towards the edge. Once fear has declined in that situation, one could repeat this on progressively higher balconies and on balconies without safety rails.

In contrast, recent suggestions based on inhibitory learning theory hold that the use of exposure hierarchies is not necessary and might even be detrimental for long-term fear reduction (Craske et al., 2008; Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014; Weisman & Rodebaugh, 2018). Instead, therapies based on inhibitory learning theory prescribe to conduct exposure in a variable way, going through the items of the hierarchy in a random order rather than in a linear fashion (Craske et al., 2014; Knowles & Olatunji, 2018). For example, having exposure on a different floor every time instead of gradually moving from lower to higher floors. In particular, random exposure might result in higher levels of variability in subjective fear and physiological arousal during treatment (e.g., Kircanski et al., 2012). Other research has demonstrated that such variability is associated with less fear responding at retest; perhaps because it facilitates memory consolidation (Brown, LeBeau,

Chat, & Craske, 2017; Cain, Blouin, & Barad, 2004; Rescorla, 2006; Culver, Stoyanova, & Craske, 2012; Kircanski et al., 2012).

In addition, it has been argued that random exposure might allow for more expectancy violation, which according to inhibitory learning theory drives fear reduction during exposure therapy (Craske et al., 2008). Expectancy violation refers to the mismatch between the expected outcome (e.g., falling off the balcony and dying) and the actual outcome (e.g., not falling off the balcony and dying; Rescorla & Wagner, 1972). The strength of this mismatch is assumed to determine the amount of (long-term) fear reduction. It should hereby be noted that the accuracy of the theorem that the net amount of prediction error will be higher throughout the course of random (versus hierarchical) exposure treatment depends on certain parameters (e.g., concerning the amount of generalization between the presented feared stimuli). In other words, theoretically speaking, this will not always be the case.

To date, two clinical studies have directly compared random exposure with a hierarchical approach. In a study by Lang and Craske (2000), participants fearful of heights had one hour of exposure divided into twelve 5-minute exercises. In each exercise, participants were asked to approach the rail at a particular height. In the hierarchical group, participants practiced from the lowest to the highest floor in a hierarchical order and always approached the rail in the same way. Participants in the random group practiced at the different heights in a random order and varied in how they approached the rail. Moreover, the random group practiced at two different locations, whereas participants in the hierarchical group always practiced at the same location. Results revealed no differences in (long-term) treatment outcomes between both groups in the fear of heights. In a second study on this topic by Kircanski et al. (2012), participants with contamination fears were assigned to a random group or a hierarchical group. In the hierarchical group, participants were exposed to one item per session, progressing from the lowest to the highest item of the participant's fear hierarchy

across the three sessions. So, they practiced with the least fear-provoking item during the first session, with the middle fear-provoking item during the second session and with the highest fear-provoking item during the third session. Within each session, the item was approached in the same gradual way and all exercises had the same duration. Participants in the random group were exposed to all items in a random order of difficulty during each session and for varying durations. Similar to the results of Lang and Craske (2000), no differences between both groups were found in the short- and long-term outcomes of exposure.

Hence, both clinical studies relevant to this topic did not find a difference in treatment outcome between random and hierarchical exposure. However, not only the order in which the feared stimuli were presented was manipulated in these studies, but also the number of stimuli and contexts as well as the timing of the exposure trials. In the current study, we aimed to compare the effects of hierarchical versus random exposure on fear responding using a fear conditioning paradigm, controlling for all other variables than the order in which the feared stimuli were presented. On the first day, participants went through an acquisition phase in which one stimulus was always paired with an electric shock (i.e., danger cue; CS+) and another stimulus was never paired with shock (i.e., safe cue; CS-). In the extinction phase, which took place on the second day, participants were presented with the CS+ and CS- and a set of morphed stimuli in between the CS+ and CS-. None of these cues was followed by shock and, crucially, we manipulated the sequence in which the stimuli were presented. Participants in the hierarchical extinction (HE) condition were first presented with the CS- and then gradually proceeded to the CS+. So, after presenting the CS-, the morph that was most similar to it was presented. Subsequently, the most similar to that one was presented, etc. In the random extinction (RE) condition the same stimuli were presented but in a random order. On the third day, we tested fear responding to the CS+, the CS- and a new third stimulus which served as a generalization stimulus (GS). In line with the assumptions of

inhibitory learning theory, we expected lower fear responding at test in the RE compared to the HE condition. Additionally, higher physiological arousal, higher levels of variability in US-expectancy ratings and higher expectancy violation were expected during extinction in the RE condition compared to the HE condition. It was predicted that, irrespective of condition, higher physiological arousal, higher variability in US-expectancy ratings and higher expectancy violation throughout extinction would be associated with lower fear responding at test.

Method

Participants

Sixty-four participants ($M_{\text{age}} = 22.31$; $SD = 6.04$; 53 females) took part in the experiment in exchange for course credit or payment. Half of the participants were assigned to the hierarchical extinction (HE) condition, the other half to the random extinction (RE) condition. Exclusion criteria were pregnancy, cardio-pulmonary conditions, psychiatric or neurological disorders (e.g., epilepsy) and wrist pain. All participants gave informed consent and the study was approved by the ethical committee of the Faculty of Psychology and Educational Sciences of the University of Leuven.

Apparatus and Stimuli

Conditioned stimuli. Two pictures of neutral male faces were used as conditioned stimuli (CS). It was counterbalanced which of the two faces functioned as the CS+ and which one as the CS-. The extinction stimuli (ES) were eight stepwise morphs between the two CSs (Figure 1; e.g., Leer et al., 2017; Lenaert et al., 2012). A third picture of another neutral male face served as the generalization test stimulus (GS; see Figure 1). All pictures were approximately 88 mm wide and 132 mm high (332×500 pixels) and were presented against a black background on a 19 inch Dell monitor (type P1911, resolution: 1440×900 at 60 Hz).

Unconditioned stimulus. As an unconditioned stimulus (US), a 2 ms electrocutaneous stimulus was administered to the participant's right wrist. The US was delivered by a Digitimer DS7A constant current stimulator (Hertfordshire, UK), through a pair of V91-01 8-mm reusable Bilaney Ag/AgCl electrodes filled with K-Y jelly.

Measures

US-expectancy ratings. US-expectancies were rated by the participant on a trial-by-trial basis on an 11-point scale. The scale ranged from 0 = "certainly no shock" to 10 = "certainly shock". Participants were instructed to click a position on the scale that matched their expectancy using the computer mouse. The rating scale appeared onscreen 200 ms after stimulus onset. Participants had 7 s to respond. After 7 s or after participants gave their rating the scale disappeared.

Skin conductance response (SCR). Electrodermal responding was measured through a pair of disposable Biopac EL 507 electrodes (contact area = 95 mm²) filled with isotonic paste and attached to the hypothenar site of the left hand palm. A Coulbourn LabLinc V Isolated Skin Conductance coupler (model V71-23, Coulbourn Instruments, Allentown, PA) applied a constant voltage of 0.5 Volts through these electrodes. The analog signal was digitized at 10 Hz by a NI PCI 3221 data acquisition card (National Instruments Corporation, Austin, Texas) from 2 s prior stimulus onset until 6 s after stimulus offset.

Matlab was used to parse the data offline in a trial-by-trial fashion and locate means and maxima relative to CS occurrence. For each trial the mean values measured at baseline (i.e., 2 s prior to stimulus onset) were subtracted from the maximum values measured in the 7.5 s after stimulus onset (Pineles, Orr, & Orr, 2009). Negative difference scores were recoded to zero. The resulting SCR amplitudes (μ Siemens) were range-corrected with each participant's largest unconditioned response measured during US present trials (Lykken &

Venables, 1971). To normalize the data, a square root transformation was subsequently applied (Dawson, Schell, & Fillion, 2000).

Expectancy violation during extinction. Since US-expectancies during extinction reflect the mismatch between the expected outcome and the actual outcome (i.e., no US), we measured expectancy violation throughout the extinction phase by calculating the sum of the US-expectancy ratings across all extinction trials for each participant.

Physiological arousal during extinction. Physiological arousal throughout extinction was measured by calculating for each participant the sum of the range-corrected, square root transformed SCR across all extinction trials.

Variability in US-expectancies during extinction. Variability throughout the extinction phase was measured by calculating the standard deviation (SD) of the US-expectancy ratings across all extinction trials for each participant (e.g., Culver et al., 2012).

Procedure

The experiment consisted of three assessments on three subsequent days (i.e., Day 1, Day 2 and Day 3).

Day 1 started with general instructions and completion of the informed consent. Subsequently, the electrodes were attached and the intensity of the US was set to a level that was described as “definitely uncomfortable and demanding some effort to tolerate, but not painful” for each individual participant using a standard work-up procedure. The mean US intensity in this study was 21.91 mA ($SD = 12.94$). Following this, instructions about the experimental task were presented on the computer screen and orally repeated by the experimenter. Specifically, participants were instructed that pictures of faces would appear on the screen and that some of these faces could be followed by an electric shock. It was stated

that participants' task was to predict the occurrence of the electric shock. Participants were then trained to use the rating scale in three practice trials. After each practice trial participants received feedback on whether or not the scale was filled in correctly. No stimuli were presented during these practice trials.

A schematic overview of the experimental phases is displayed in Table 1. On Day 1, the experimental task started with a pre-exposure phase in which the CS+ and CS- were presented twice without the US. This phase was added to reduce orienting responses in the skin conductance measure. The subsequent acquisition phase consisted of eight presentations of the CS+ followed by the US and eight CS- trials in absence of the US. Trial order was semi-randomized with no more than two consecutive trials with the same stimulus. To control for order-effects in both the pre-exposure phase and the acquisition phase, the CS+ was presented first for half of the participants in each condition and the CS- was presented first for the other half (e.g., Lovibond, 2003; Vervliet et al., 2005).

On Day 2 participants returned to the laboratory to continue with the extinction training. After attachment of the SCR and shock electrodes, participants were instructed that their task was the same as on Day 1 and to keep in mind what they had previously learned about the faces. In both groups the same set of stimuli was presented, but the sequence in which this happened was manipulated. In the HE condition, participants were first presented with the CS- and then gradually moved from the CS- to the CS+ along the morphs. Each stimulus displayed in Figure 1 was presented twice, resulting in a total of 20 extinction trials (i.e., CS- → CS- → ES8 → ES8 → ES7 → ES7 → ES6 → ES6 → ...). Participants in the RE condition were presented with the same stimuli but in a random order (e.g., ES6 → ES2 → ES7 → ES6 → CS- → ES3 → CS+ → ES8 → ...). In this condition the stimulus sequence was determined randomly per individual participant. No USs were administered during extinction training.

On Day 3, the SCR and shock electrodes were again attached. Participants were told that their task was the same as previously and that they could use the information that they had previously learned. Subsequently, the CS+ and CS- were presented (in counterbalanced order), followed by the GS. This test cycle was repeated four times.

In all experimental phases, each trial started with a 2 s blank screen to measure baseline skin conductance. Stimuli were presented for 8 s. If applicable, USs were delivered 7.5 s after stimulus onset. Inter-trial intervals (ITI) were set at an average of 8 s (range 7-9 s). Trial order, stimulus presentation, ITI and registration of the dependent variables were controlled by Affect 4.0 (Spruyt, Clarysse, Vansteenwegen, Baeyens, & Hermans, 2010).

Results

US-expectancy ratings

Figure 2 displays the mean US-expectancy ratings for the CSs, ESs and GS in the HE and RE condition. The first and last acquisition trials, 20 extinction trials, and four test trials are presented. A Greenhouse-Geisser correction was applied when the assumption of sphericity was violated.

Acquisition phase. The left panel of Figure 2 suggests that participants learned to discriminate between the CS+ and CS- from the first to the last acquisition trial, with no differences between conditions. This was confirmed by a repeated measures Analysis of Variance (rmANOVA) with Trial (2 levels: acq1, acq8) and Stimulus (2 levels: CS+, CS-) as within-subjects factors and Condition (2 levels: HE, RE) as between-subjects factor. A main effect of Trial, $F(1, 62) = 121.41, p < .001, \eta^2_p = 0.66$, a main effect of Stimulus, $F(1, 62) = 524.95, p < .001, \eta^2_p = 0.89$, and a significant Trial \times Stimulus interaction, $F(1, 62) = 821.83, p < .001, \eta^2_p = 0.93$, were found. The Trial \times Stimulus interaction confirms the development

of CS+/CS- discrimination throughout the acquisition phase. As expected in this phase of the experiment, no significant Trial \times Stimulus \times Condition interaction was found, indicating that acquisition learning did not differ between the HE and RE condition. For reasons of completeness, we also report the main effect of Condition, $F(1, 62) = 0.06, p = .80, \eta^2_p = 0$, the Trial \times Condition interaction, $F(1, 62) = 0, p = 1, \eta^2_p = 0$, and the Stimulus \times Condition interaction, $F(1, 62) = 4.47, p < .05, \eta^2_p = 0.07$.

Extinction phase. The middle panel of Figure 2 indicates a different course of the extinction phase in the RE and HE condition. A rmANOVA with Trial (20 levels: ext1 to ext20) as within-subjects factor and Condition (2 levels: HE, RE) as between-subjects factor revealed a significant Trial \times Condition interaction, $F(19, 1178) = 16.83, p < .001, \eta^2_p = 0.21$, confirming a different course of the extinction phase in both conditions. For the sake of completeness, we also report the main effect of Trial, $F(19, 1178) = 2.23, p < .01, \eta^2_p = 0.035$, and the main effect of Condition, $F(1, 62) = 7.15, p = .01, \eta^2_p = 0.21$.

Test phase. The right panel of Figure 2 suggests no differences in CS+/CS- discrimination between the RE and HE condition on the first test trial. Evidence for this was provided by a rmANOVA with Stimulus (2 levels: CS+, CS-) as a within-subjects factor and Condition (2 levels: HE, RE) as a between-subjects factor. A significant main effect of Stimulus, $F(1, 62) = 106.08, p < .001, \eta^2_p = 0.63$, indicates that there was discrimination between the CS+ and CS- on the first test trial. However, no Stimulus \times Condition interaction was found, $F(1, 62) = 0.01, p = .94, \eta^2_p = 0$. This result suggests that the data provide no evidence for a difference between the RE and HE condition in CS+/CS- discrimination on the first test trial. The main effect of Condition was not significant, $F(1, 62) = 0.12, p = .94, \eta^2_p = 0$.

Similar results were found with regard to the GS/CS- discrimination on the first test trial. A rmANOVA on the first test trial with Stimulus (2 levels: GS, CS-) as a within-subjects factor and Condition (2 levels: HE, RE) as a between-subjects factor revealed a significant

main effect of Stimulus, $F(1, 62) = 5.55, p < .05, \eta^2_p = 0.08$, but no Stimulus \times Condition interaction, $F(1, 62) = 0.45, p = .50, \eta^2_p = 0.01$. For reasons of completeness, we also report the main effect of Condition, $F(1, 62) = 0.05, p = .84, \eta^2_p = 0$.

Additionally, we tested whether the course of the test phase differed between the HE and RE condition by conducting a rmANOVA with Stimulus (2 levels: CS+, CS-) and Trial (4 levels: Test1, Test2, Test3, Test4) as within-subjects factors and Condition (2 levels: HE, RE) as a between-subjects factor. A main effect of Trial, $F(3, 186) = 59.21, p < .001, \eta^2_p = 0.49$, and a main effect of Stimulus, $F(1, 62) = 89.39, p < .001, \eta^2_p = 0.59$, were found. The Trial \times Stimulus interaction was significant, $F(3, 186) = 19.34, p < .001, \eta^2_p = 0.24$, indicating that there was re-extinction over the course of the test phase. However, the Trial \times Stimulus \times Condition interaction was not significant, $F(3, 186) = 0.06, p = .98, \eta^2_p = 0.01$, suggesting that the data do not provide evidence for a different course of the test phase in both conditions. For the sake of completeness, we also report the main effect of Condition, $F(1, 62) = 0.34, p = .56, \eta^2_p = 0.01$, the Trial \times Condition interaction, $F(3, 186) = 0.05, p = .84, \eta^2_p = 0$, and the Stimulus \times Condition interaction, $F(3, 186) = 0.03, p = .87, \eta^2_p = 0$.

Skin-conductance response (SCR)

Figure 3 displays the range-corrected, square root transformed SCR for the CSs, ESs and GS in the HE and VE condition. The first and last acquisition trials, 20 extinction trials, and four test trials are displayed. A Greenhouse-Geisser correction was applied when the assumption of sphericity was violated.

Acquisition phase. A rmANOVA with Trial (2 levels: acq1, acq8) and Stimulus (2 levels: CS+, CS-) as within-subjects factors and Condition (2 levels: HE, RE) as between-subjects factor revealed no main effect of Trial, $F(1, 62) = 0.25, p = .62, \eta^2_p = 0.01$, a significant main effect of Stimulus, $F(1, 62) = 4.25, p < .05, \eta^2_p = 0.06$, and a significant Trial \times Stimulus

interaction, $F(1, 62) = 4.08, p < .05, \eta^2_p = 0.06$. The significant Trial \times Stimulus interaction suggests that participants started to discriminate between the CS+ and CS- over the course of acquisition. No significant Trial \times Stimulus \times Condition interaction was found, $F(1, 62) = 0.62, p = .44, \eta^2_p = 0.01$. As expected, the main effect of Condition, $F(1, 62) = 2.97, p = .09, \eta^2_p = 0.05$, the Trial \times Condition interaction, $F(1, 62) = 0.30, p = .59, \eta^2_p = 0.01$, and the Stimulus \times Condition interaction, $F(1, 62) = 0, p = .99, \eta^2_p = 0$, were not significant.

Extinction phase. To test for group differences in the course of extinction, a rmANOVA with Trial (20 levels: ext1 to ext20) as within-subjects factor and Condition (2 levels: HE, RE) as between-subjects factor was performed. This resulted in a significant Trial \times Condition interaction, $F(19, 1178) = 2.08, p < .05, \eta^2_p = 0.03$, suggesting that the RE and HE condition differed in the course of extinction. For reasons of completeness, we also report the main effect of Trial, $F(19, 1178) = 1.56, p = .13, \eta^2_p = 0.025$, and the main effect of Condition, $F(1, 62) = 4.31, p < .05, \eta^2_p = 0.07$.

Test phase. Group differences in fear responding on the first trial of the test phase were tested by a rmANOVA with Stimulus (2 levels: CS+, CS-) as a within-subjects factor and Condition (2 levels: HE, RE) as a between-subjects factor. This analysis revealed a main effect of Stimulus, $F(1, 62) = 4.74, p < .05, \eta^2_p = 0.07$, but no significant Stimulus \times Condition interaction, $F(1, 62) = 2.59, p = .11, \eta^2_p = 0.04$. As such, the data provide no evidence for differences between the RE and HE condition with regard to the CS+/CS- discrimination on the first test trial. The main effect of Condition was not significant, $F(1, 62) = 0.001, p = .97, \eta^2_p = 0$.

In addition, with regard to the GS, a rmANOVA with Stimulus (2 levels: GS, CS-) as a within-subjects factor and Condition (2 levels: HE, RE) as a between-subjects factor revealed no significant main effect of Stimulus, $F(1, 62) = 0.17, p = .68, \eta^2_p = 0.01$, and no significant Stimulus \times Condition interaction, $F(1, 62) = 1.81, p = .18, \eta^2_p = 0.03$. These results

suggest that the data contain no evidence for a GS/CS- discrimination on the first test trial in SCR or between-condition differences therein. For reasons of completeness, we also report the main effect of Condition, $F(1, 62) = 0.12, p = .73, \eta^2_p = 0$.

Differences between the RE and HE condition in the course of the test phase were examined by conducting a rmANOVA with Stimulus (2 levels: CS+, CS-) and Trial (4 levels: Test1, Test2, Test3, Test4) as within-subjects factors and Condition (2 levels: HE, RE) as a between-subjects factor. This revealed a main effect of Trial, $F(3, 186) = 15.01, p < .001, \eta^2_p = 0.20$, no main effect of Stimulus, $F(1, 62) = 2.74, p = .10, \eta^2_p = 0.04$, no significant Trial \times Stimulus interaction, $F(3, 186) = 1.47, p = .23, \eta^2_p = 0.02$, and no significant Trial \times Stimulus \times Condition interaction, $F(3, 186) = 1.43, p = .24, \eta^2_p = 0.02$. Hence, we found no evidence for differences between the RE and HE condition in the course of the test phase. The main effect of Condition, $F(1, 62) = 1.06, p = .31, \eta^2_p = 0.02$, the Trial \times Condition interaction, $F(3, 186) = 1.02, p = .38, \eta^2_p = 0.02$, and the Stimulus \times Condition interaction, $F(1, 62) = 0.30, p = .59, \eta^2_p = 0.01$, were not significant.

Expectancy violation, physiological arousal, variability throughout extinction and their correlation with fear responding at test

It was predicted that random extinction would result in higher *expectancy violation* throughout the extinction phase compared to hierarchical extinction. Results of an independent-samples t-test confirm that the sum of the US-expectancy ratings was significantly higher in the RE condition ($M = 61.69; SD = 5.16$) compared to the HE condition ($M = 41.16; SD = 5.70$), $t(62) = -2.67, p = .01$.

In line with inhibitory learning theory, we predicted that individuals with higher expectancy violation during extinction would show less fear responding (i.e., CS+/CS- discrimination) on the first test trial. A significant positive correlation ($r = .27, p = .03$) was

found between CS+/CS- discrimination on the first test trial and the sum of US-expectancies during extinction, suggesting that the higher the expectancy violation, the more discriminative fear responding during test. No significant correlation was found between the sum of the US-expectancies during extinction and differential SCR on the first test trial.

Furthermore, higher *physiological arousal* was expected throughout extinction in the RE condition compared to the HE condition. An independent-samples t-test, $t(62) = -2.08, p = .04$, confirmed that the RE condition ($M = 18.48; SD = 1.57$) showed higher SCR throughout extinction compared to the HE condition ($M = 14.04; SD = 1.46$). Physiological arousal during extinction was, however, not associated with fear responding on the first test trial. No significant correlations were found between the sum of SCR across all extinction trials and discriminative fear responding on the first test trial in both US-expectancies and SCR.

Finally, we tested whether participants in the RE condition reported higher *variability* than those in the HE condition. An independent-samples t-test confirms that participants in the RE condition showed higher variability in their US-expectancy ratings ($M = 2.81; SD = 0.12$) than participants in the HE condition ($M = 2.22; SD = 0.23$), $t(62) = -2.25, p = .03$. In addition, a significant positive correlation ($r = .65, p < .001$) was found between individuals' SD in US-expectancies during extinction and CS+/CS- discrimination in the US-expectancies on the first test trial. This result suggests that higher levels of variability during extinction are associated with higher discriminative fear responding at test. The degree of variability in US-expectancy ratings during extinction was not correlated with differential SCR on the first test trial. The correlations between expectancy violation, physiological arousal, variability during extinction and fear responding to the CS+ on the first test trial lead to the same conclusions as the correlations with CS+/CS- discrimination. Notably, the correlation between physiological arousal and SCR on the first test trial was significantly positively correlated with CS+ ($r = .46, p < .001$), whereas the correlation with CS+/CS- discrimination was not significant.

However, since the direction of this correlation was opposite to what is predicted by inhibitory learning theory, this did not influence the main conclusions.

Moderation of the relation between expectancy violation, variability throughout extinction and fear responding at test

In sum, we did not find differences between both conditions in CS+/CS- discrimination at test, although (1) an independent-samples t-test showed that the sum of US-expectancy ratings as well as variability throughout extinction were higher in the RE condition than in the HE condition and (2) the sum of the US-expectancies and variability throughout extinction were correlated with CS+/CS- discrimination at test in the entire sample. To assess this further we followed this up with a moderation analysis. This analysis showed that the correlation between expectancy violation and CS+/CS- discrimination on the first test trial was moderated by condition ($\Delta R^2 = .085$, $\Delta F(1, 60) = 6.11$, $p = .02$, $\beta = 0.59$, $t(60) = 2.47$, $p = .02$). In the HE condition higher expectancy violation predicted more CS+/CS- discrimination on the first test trial, whereas expectancy violation was not predictive for CS+/CS- discrimination on the first test trial in the RE condition. Importantly, although the correlations differ between both conditions, both results are not in line with the predictions of inhibitory learning theory (i.e., that more expectancy violation would predict lower CS+/CS- discrimination at test). A second moderation analysis was performed with regard to the relation between variability during extinction and CS+/CS- discrimination in US-expectancies on the first test trial. This analysis showed that condition did not moderate the relation between variability during extinction and discriminative fear responding on the first test trial in US-expectancies ($\Delta R^2 = .002$, $\Delta F(1, 60) = 0.27$, $p = .61$, $\beta = 0.18$, $t(60) = 0.52$, $p = .61$).

Discussion

Exposure therapy can either start with little fear-eliciting stimuli and gradually progress to more distressing stimuli (i.e., hierarchical approach) or the items can be presented in a random order. In the current study a procedure was developed to investigate the effects of hierarchical versus random extinction on fear responding using a fear conditioning procedure. On Day 1 participants learned that one stimulus (i.e., CS+) was always followed by an electric shock, whereas another stimulus (i.e., CS-) was never followed by shock. During extinction training on Day 2, participants were presented with the CS+, CS-, and eight morphed stimuli between the CS+ and CS-. In the hierarchical extinction (HE) condition, we aimed to model hierarchical exposure by first presenting participants with the CS-, followed by the morph that was most similar to it, subsequently the most similar to that one, etc. Participants in the random exposure (RE) condition were presented with the same set of stimuli in a random order. On Day 3, we tested for fear responding to the CS+, CS- and a third generalization stimulus (GS). No shocks were administered on Day 2 and 3. In line with the assumptions of the inhibitory learning theory, higher fear responding at test was expected in the HE compared to the RE condition. In addition, higher expectancy violation, higher physiological arousal and higher variability were expected during extinction in the RE condition compared to the HE condition

However, the results provided no evidence for differences between the HE and RE condition in differential fear responding on the first test trial, as measured by US-expectancy ratings and SCR. In addition, we did not find evidence for differences in the course of the test phase between both conditions. As predicted, higher expectancy violation, higher physiological arousal and higher levels of variability in US-expectancy ratings were found throughout extinction in the RE condition compared to the HE condition. However, physiological arousal during extinction was not associated with fear responding during test in the entire sample. Opposite to what was predicted, higher levels of variability in US-

expectancy ratings and higher expectancy violation during extinction predicted higher differential fear responding at test as evidenced by correlations calculated for the complete sample. Moderation analyses demonstrated that the correlations in the individual conditions were not in line with the predictions of inhibitory learning theory either.

These findings are not in line with the predictions made by the inhibitory learning theory. Although random extinction successfully induced higher levels of expectancy violation and physiological arousal, we did not find the predicted effects on subsequent fear responding. Moreover, the observation that higher levels of variability during extinction predicted higher subsequent fear responding is opposite to previous findings by Culver et al. (2012), Kircanski et al. (2012), and Brown et al. (2017). They found that higher variability throughout exposure was associated with lower fear responding at follow-up. Notably, in the study of Brown et al. (2017) variability in US-expectancies during extinction was operationalized by calculating the sum of the absolute value of differences in US-expectancies on each consecutive trial. Similar analyses on our data did, however, not change our conclusions. Whereas Brown et al. (2017), similar to the current study, used variability in US-expectancies to test this hypothesis, Culver et al. (2012) and Kircanski et al. (2012) focused on variability in *subjective fear ratings*. Arguably, as compared to subjective fear ratings, US-expectancy ratings allow for a more direct measure of the expectancy violation that inhibitory learning theory deems important. One can indeed assume that US-expectancies during extinction reflect expectancy violation given that the US was not administered in this phase (e.g., Sevenster, Beckers, & Kindt, 2013). However, to gauge expectancy violation even more directly, it is recommended for future research to ask participants after each US omission during extinction about their subjective feeling of “surprise.”

The predictions about random versus hierarchical extinction made by inhibitory learning theory are opposite to what recent work on updating of fear memories predicts (e.g.,

Gershman et al., 2013). In this line of research, it is assumed that large expectancy violation results in the formation of a new competing (inhibitory) memory, because the abrupt transition from acquisition to extinction indicates a novel *state* prompting the formation of such new memory (e.g., Gershman, Blei, & Niv, 2010). This approach states that if the transition between acquisition and extinction occurs more gradually, expectancy violation (or prediction error) is kept low and instead of the formation of a new memory, new information can be integrated in the original fear memory. Gershman and colleagues (2013) tested this hypothesis in rodents by manipulating the reinforcement schedule to keep the transition between acquisition and extinction gradual and expectancy violation low. They found that gradually rather than abruptly reducing the number of USs during extinction prevented the recovery of fear. In the current study, the transition between acquisition and extinction was made gradual in the HE condition, since we first presented the CS-, which evoked little expectancy violation because it was never paired by shock. Subsequently the morph most similar to the CS- was presented, followed by the morph most similar to that stimulus, etc. This way, the differences between the subsequent stimuli were very small and we moved to the CS+ in a very gradual way, arguably keeping expectancy violation low. However, our findings are not in line with this approach either, because we did not find evidence for hierarchical extinction resulting in less fear responding during test than random extinction. Notably, we found overall lower expectancy violation in the HE condition compared to the RE condition, but nevertheless US-expectancies increased in the HE condition by the end of extinction training, when being presented with stimuli that were more similar to the CS+ and the CS+ itself. As predicted, the hierarchical approach thus resulted in less expectancy violation relative to the random approach during extinction, but still elicited a certain level of expectancy violation. The same holds for variability and sustained physiological arousal

during extinction: There was still variability and physiological arousal in the HE condition, although less as compared to the RE condition.

In theoretical approaches such as the inhibitory learning theory, random exposure is seen as one of several possible strategies to induce variability throughout exposure (Craske et al., 2008). The effects of random exposure have typically been tested in combination with other strategies to induce variability, such as the use of multiple stimuli and contexts or varying the timing of exposure trials. So far, two clinical studies directly compared the effects of random versus hierarchical exposure therapy (i.e., Kircanski et al., 2012; Lang & Craske, 2000; see introduction). Despite the use of multiple strategies to induce variability, both studies found no evidence for random exposure being more effective than hierarchical exposure. These results are in line with the results of the current study.

It can be considered a unique strength of the current study that both conditions only differed in the order in which the feared stimuli were presented, while controlling for all other variables. The current study's focus on internal validity might, however, be at the expense of its external validity. As inherent to this type of laboratory treatment research, clinical exposure therapy is reduced to its putative core mechanisms (Scheveneels, Boddez, Vervliet, & Hermans, 2016). We therefore recommend to test the outcomes of hierarchical versus random exposure in clinical samples, while controlling for other variables than stimulus order, in future research.

The findings of the present study suggest that other factors should come into play when deciding to choose one approach or the other, such as drop-out rates. In particular, it could be an advantage of the hierarchical approach of exposure that starting with less fear-eliciting stimuli will prevent clients from dropping out and increase the chance that they will engage in exposure to stimuli higher in the fear hierarchy. In the study of Lang and Craske (2000) higher peak fear levels were found in the random group, suggesting that random exposure

might be less feasible for clients and associated with higher drop-out from treatment. The current study does not provide an answer to this question, but it might be an interesting topic for future research.

In conclusion, we developed a novel procedure to investigate the effects of hierarchical versus random extinction in the laboratory, using a set of morphed stimuli during extinction. Arguably, this procedure is a better model for clinical exposure therapy than regular extinction training to the CS+. Clinical exposure is typically not limited to the CS+ and often the CS+ is unavailable or unidentified. The observation that presenting the stimuli in a random order resulted in higher expectancies, higher physiological arousal and higher variability than presenting them in a hierarchical order, adds to the validity of this procedure. However, the results of the current study do not provide evidence for the superiority of random extinction on long-term fear reduction.

References

- Brown, L. A., LeBeau, R. T., Chat, K. Y., & Craske, M. G. (2017). Associative learning versus fear habituation as predictors of long-term extinction retention. *Cognition & Emotion, 31*, 687-698. doi:10.1080/02699931.2016.1158695
- Cain, C. K., Blouin, A. M., & Barad, M. (2004). Adrenergic transmission facilitates extinction of conditional fear in mice. *Learning and Memory, 11*, 179-187. doi:10.1101/lm.71504
- Choy, Y., Fyer, A. J., & Lipsitz, J. D. (2007). Treatment of specific phobia in adults. *Clinical Psychology Review, 27*, 266-286. doi:10.1016/j.cpr.2006.10.002
- Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, N., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy, 46*, 5-27. doi:10.1016/j.brat.2007.10.003
- Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach, *Behaviour Research and Therapy, 58*, 10-23. doi:10.1016/j.brat.2014.04.006
- Culver, N., Stoyanova, M. S., & Craske, M. G. (2012). Emotional variability and sustained arousal during exposure. *Journal of Behavior Therapy and Experimental Psychiatry, 42*, 787-793. doi:10.1016/j.jbtep.2011.10.009
- Dawson, M. E., Schell, A. M., & Fillion, D. L. (2000). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 200-224). Cambridge: Cambridge University Press.
- Foa, E. B., Huppert, J. D., & Cahill, S. P. (2006). Emotional processing theory: An update. In B. O. Rothbaum (Ed.), *Pathological anxiety: Emotional processing in etiology and treatment* (pp. 3-24). New York: Guilford Press.
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99*, 20-35.

- Foa, E. B., & McLean, C. P. (2016). The efficacy of exposure therapy for anxiety-related disorders and its underlying mechanisms: The case of OCD and PTSD. *Annual Review of Clinical Psychology*, 12, 1-28. doi:10.1146/annurev-clinpsy-021815-093533
- Gershman, S. J., Blei, D. M., & Niv, Y. (2010). Context, learning, and extinction. *Psychological Review*, 117, 197-209. doi:10.1037/a0017808
- Gershman, S. J., Jones, C. E., Norman, K. A., Monfils, M-H., Niv, Y. (2013). Gradual extinction prevents the return of fear: Implications for the discovery of state. *Frontiers in Behavioral Neuroscience*, 7, 164. doi:103389/fnbeh.2013.00164
- Hovland, C. I. (1937). The generalization of conditioned responses. III. Extinction, spontaneous recovery, and disinhibition of conditioned and generalized responses. *Journal of Experimental Psychology*, 21, 47-62. doi:10.1037/h0055714
- Kircanski, K., Mortazavi, A., Castriotta, N., Baker, A. S., Mystkowski, J. L., Yi, R., & Craske, M. G. (2012). Challenges to the traditional exposure paradigm: Variability in exposure therapy for contamination fears. *Journal of Behavior Therapy and Experimental Psychiatry*, 43, 745-751. doi:10.1016/j.jbtep.2011.10.010
- Knowles, K. A., & Olatunji, B. O. (2018). Enhancing inhibitory learning: The utility of variability in exposure. *Cognitive and Behavioral Practice*. doi:10.1016/j.cbpra.2017.12.001
- Lang, A. J., & Craske, M. G. (2000). Manipulations of exposure-based therapy to reduce return of fear: A replication. *Behaviour Research and Therapy*, 38, 1-12. doi:10.1016/S0005-7967(99)00031-5
- Leer, A., Engelhard, I. M., Lenaert, B., Struyf, D., Vervliet, B., & Hermans, D. (2017). Eye movement during recall reduces objective memory performance: An extended replication. *Behaviour Research and Therapy*, 92, 94-105. doi:10.1016/j.brat.2017.03.002

- Lenaert, B., Claes, S., Raes, F., Boddez, Y., Joos, E., Vervliet, B., & Hermans, D. (2012). Generalization of conditioned responding: Effects of autobiographical memory specificity. *Journal of Behavior Therapy and Experimental Psychiatry*, 43, 60-66. doi:10.1016/j.btep.2010.12.010
- Lovibond, P. F. (2003). Causal beliefs and conditioned responses: retrospective revaluation induced by experience and instruction. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29, 97-106. doi:10.1037/0278-7393.29.1.97
- Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*, 8, 656-672.
- Myers, K. M., & Davis, M. (2007). Mechanisms of fear extinction. *Molecular Psychiatry*, 12, 120-150. doi:10.1038/sj.mp.4001939
- Pappens, M., Schroijen, M., Van den Bergh, O., & Van Diest, I. (2015). Retention of perceptual generalization of fear extinction. *International Journal of Psychophysiology*, 98, 520-528. doi:10.1016/j.ijpsycho.2015.01.007
- Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear condition paradigm with a long-duration conditioned stimulus. *Psychophysiology*, 46, 984-995. doi:10.1111/j.1469-8986.2009.00852
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp 64-99). New York: Appleton-Century-Crofts.
- Sánchez-Meca, J., Rosa-Alcázar, A. I., Marín-Martínez, F., & Gómez-Conesa, A. (2010). Psychological treatment of panic disorder with or without agoraphobia: A meta-analysis. *Clinical Psychology Review*, 30, 37-50. doi:10.1016/j.cpr.2009.08.011

- Scheveneels, S., Boddez, Y., Vervliet, B., & Hermans, D. (2016). The validity of laboratory-based treatment research: Bridging the gap between fear extinction and exposure treatment. *Behaviour Research and Therapy*, 86, 87-94.
doi:10.1016/j.brat.2016.08.015
- Sevenster, D., Beckers, T., & Kindt, M. (2013). Prediction error governs pharmacologically induced amnesia for learned fear. *Science*, 339, 830-833.
- Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0: A free software package for implementing psychological and psychophysiological experiments. *Experimental Psychology*, 57, 36-45. doi:10.1027/1618-3169/a000005
- Vervliet, B., Vansteenwegen, D., Baeyens, F., Hermans, D., & Eelen, P. (2005). Return of fear in a human differential conditioning paradigm caused by a stimulus change after extinction. *Behaviour Research and Therapy*, 43, 357-371.
doi:10.1016/j.brat.2004.02.005
- Weisman, J. S., & Rodebaugh, T. L. (2018). Exposure therapy augmentation: A review and extension of techniques informed by an inhibitory learning approach. *Clinical Psychology Review*, 59, 41-51. doi:10.1016/j.cpr.2017.10.010

Tables and Figures

Table 1

Overview of the experimental phases

Day 1		Day 2		Day3
	Pre exposure	Acquisition	Extinction	Test
HE condition	CS+ (2)	CS+ (8)	CS- (2) → ES8 (2) → ES7 (2) → ES6	CS+ (4)
	CS- (2)	CS- (8)	(2) → ES5 (2) → ES4 (2) → ES3 (2)	CS- (4)
			→ ES2 (2) → ES1 (2) → CS+ (2)	GS (4)
RE condition	CS+ (2)	CS+ (8)	Random per participant: CS- (2); ES8	CS+ (4)
	CS- (2)	CS- (8)	(2); ES7 (2); ES6 (2); ES5 (2); ES4	CS- (4)
			(2); ES3 (2); ES2 (2); ES1 (2); CS+ (2)	GS (4)

Note. HE refers to hierarchical extinction, RE refers to random extinction. CSs, ESs (counterbalanced) and the GS are pictures of human faces. During acquisition ‘+’ refers to the administration of the US and ‘-’ to the absence of the US. During the other phases no USs are administered. The number of trials is indicated between parentheses.



Figure 1. CS+, CS-, eight morphs that served as extinction stimuli (ES) and the generalization test stimulus (GS). CS+, CS- and ESs were counterbalanced.

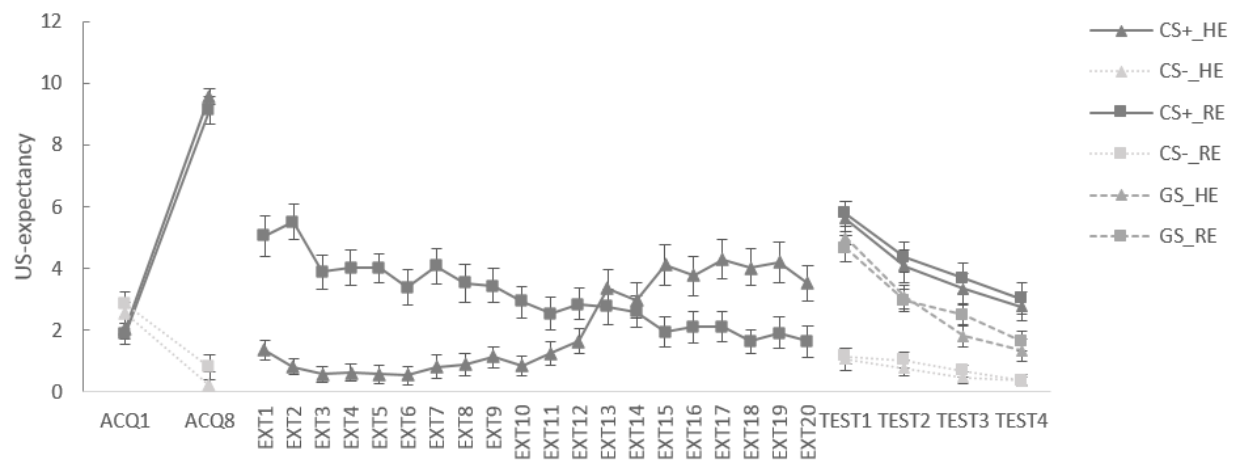


Figure 2. Mean US-expectancy ratings for the first and last acquisition trial, all extinction trials and all test trials per condition and CSs/GS. Error bars represent standard error of the means.

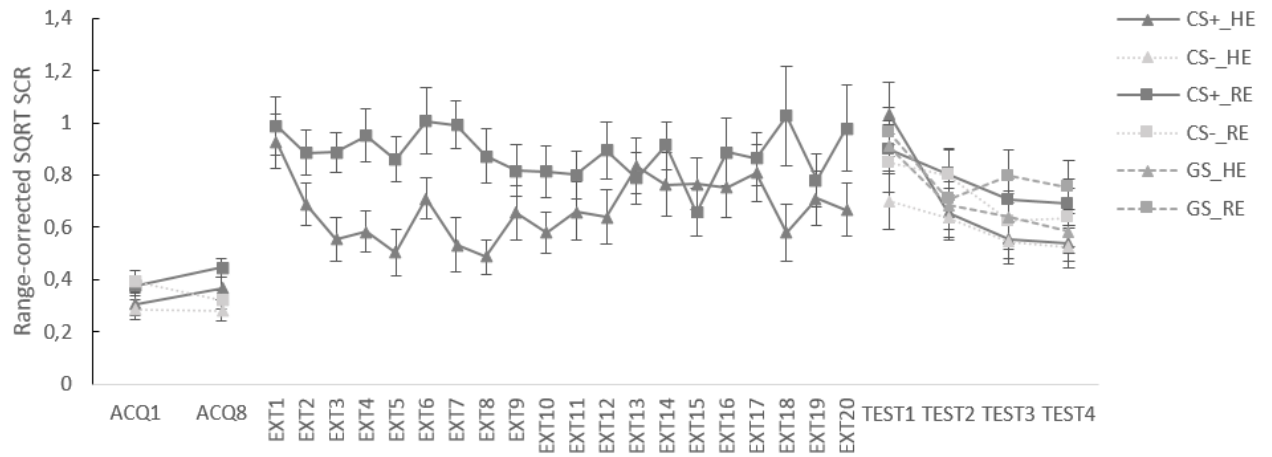


Figure 3. Range-corrected SQRT SCR for the first and last acquisition trial, all extinction trials and all test trials per condition and CS/GS. Error bars represent standard error of the means.